Lecture 6

Recall:

Thm (Portmanteau): For any $g: Z \to \mathbb{R} \cup \{+\infty\}$ lsc, bdd below, the functional

$$\underset{\in \mathcal{P}(Z)}{\mu} \mapsto \underset{\in \mathbb{R}\cup\{+\infty\}}{\int_Z g \, d\mu}$$

is lsc wrt narrow topology.

Cor: Given metric spaces (X, d_X) and (Y, d_Y), for any function $c: X \times Y \to \mathbb{R} \cup \{+\infty\}$ that is lsc and bdd below, $\mathbb{K}_c(\gamma)$ is lsc in the narrow topology.

Combining these compactness and lsc results, by the direct method of the calculus of variations...

<u>Thm</u>: Given Polish spaces (X, d_X) and (Y, d_Y), for any function $c: X \times Y \to \mathbb{R} \cup \{+\infty\}$ that is lsc and bdd below and for any $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$, there exists $\gamma_* \in \Gamma(\mu, \nu)$ satisfying

$$\gamma_* = \min_{\gamma : \gamma \in \Gamma(\mu,\nu)} \int c(x,y)\, d\gamma(x,y).$$

A solution of Kantorovich's problem exists!

Why was the narrow topology the "right" topology?

Let $(X, \|\cdot\|_X)$ be a normed vector space

n.v.s.

Let $(X^*, \|\cdot\|_{X^*})$ be its dual space, that is, the set of all bdd linear functionals on X with

$$\|y\|_{X^*} = \sup_{x \in X, \|x\| \leq 1} y(x)$$

Given $x \in X$, $y \in X^*$ let

$$\langle y, x \rangle := y(x).$$

Topologies on n.v.s.

$\boxed{\text{on } X}$

- norm, $\|x_n - x\|_X \to 0$
- weak, $\underbrace{\langle y, x - x_n \rangle}_{y(x) - y(x_n)} \to 0 \;\; \forall\, y \in X^*$

$\boxed{\text{on } X^*}$

- norm, $\|y_n - y\|_{X^*} \to 0$
- weak, $\langle z, y_n - y \rangle \to 0 \;\; \forall\, z \in (X^*)^*$
- weak-*, $\langle y_n - y, x \rangle \to 0 \;\; \forall\, x \in X$

$\mathcal{X}$ is a locally compact metric space

<u>Topologies on $\mathcal{M}_s(\mathcal{X})$</u>

NOT METRIZABLE

- strong, $\mu_n(A) \to \mu(A) \quad \forall A \in \mathcal{B}(\mathcal{X})$
- total variation norm

$$\|\mu - \nu\|_{TV} = \sup_{A \in \mathcal{B}(\mathcal{X})} |\mu(A) - \nu(A)|$$

$$= \sup_{\substack{f \in C_0(\mathcal{X}) \\ \|f\|_\infty \leq 1}} \int f \, d\mu - \int f \, d\nu$$

...where...

$$C_0(\mathcal{X}) = \{ f : \mathcal{X} \to \mathbb{R} : \forall \varepsilon > 0, K_\varepsilon \subset\subset \mathcal{X} \text{ s.t. } |f(x)| \leq \varepsilon \ \forall x \in K_\varepsilon^c \}$$

Banach space	$(C_0(\mathcal{X}), \|\cdot\|_\infty)$	$(C_b(\mathcal{X}), \|\cdot\|_\infty)$
Dual space	$(\mathcal{M}_s(\mathcal{X}), \|\cdot\|_{TV})$ $\Downarrow$	:(
Weak-* topology	wide topology $\Leftarrow$	narrow topology

$$\mu_n \xrightarrow{\text{wide}} \mu \iff \int f \, d\mu_n \to \int f \, d\mu \quad \forall f \in C_0(\mathcal{X})$$

Which topology for Kantorovich's problem?

Exercise 10: Give an example of $\mu \in P(\mathbb{R})$ for which $\Gamma(\mu,\mu)$ is not compact wrt $\|\cdot\|_{TV}$

Similarly, wide topology is too weak to ensure lsc of $\mathbb{K}_c(\gamma)$.

Moral: Wide convergence can allow mass to "escape to $+\infty$".
OTOH, given a narrowly convergent sequence of probability measures, the limit <u>must</u> be a prob measure.

However...

Prop: Given a locally compact Polish space X and $\{\mu_n\}_{n=1}^{\infty} \subseteq P(X)$, $\mu \in P(X)$, then

$$\mu_n \to \mu \text{ narrowly}$$

if and only if

$$(*) \begin{cases} \lim_{n\to\infty} \int f \, d\mu_n = \int f \, d\mu \quad \forall f \in C_c(X). \end{cases}$$

In particular,

$$\left.\begin{array}{l} \text{wide convergence} \\ \qquad + \\ \text{conservation of} \\ \text{mass \& positivity} \end{array}\right\} \Leftrightarrow \text{narrow conv}$$

Pf: Suppose (*) holds.
Fix $f \in C_b(X)$ arbitrary.

Let B_k denote the closed ball centered at $x_0 \in X$ with radius k.

Cutoff functions on $[0,+\infty)$...
Define $\varphi_k, \xi_k : [0,+\infty) \to [0,1]$ cts via...

$\boxed{k=1}$

- $\varphi_1 \equiv 1$ on $[0,1]$, $\varphi \equiv 0$ on $[2,+\infty)$

$\boxed{k>1}$

- $\xi_k \equiv 1$ on $[0,k]$, $\xi_k \equiv 0$ on $[k+1,+\infty)$
- $\varphi_k := \max\{\varphi_{k-1}, \xi_k\}$

so $\varphi_k \equiv 1$ on $[0,k]$, $\varphi_k \equiv 0$ on $[k+1,+\infty)$
and $\varphi_k \nearrow 1$ as $k \to +\infty$.

Cutoff functions on $\mathcal{X}$...

Define $\eta_k : \mathcal{X} \to [0,1]$ by
$$\eta_k(x) := \varphi_k(d(x, x_0)).$$

Then η_k cts, $\eta_k \equiv 1$ on B_k, $\eta_k \equiv 0$ on B_{k+1}^c, and $\eta_k \nearrow 1$.

Note that, $\forall\, k \in \mathbb{N}$,

$$\liminf_{n\to\infty} \int f + \|f\|_\infty \, d\mu_n \geq \liminf_{n\to\infty} \int (f + \|f\|_\infty)\eta_k \, d\mu_n$$
$$\text{I} \qquad \overset{(*)}{=} \int (f + \|f\|_\infty)\eta_k \, d\mu$$

$$\limsup_{n\to\infty} \int f - \|f\|_\infty \, d\mu_n \leq \limsup_{n\to\infty} \int (f - \|f\|_\infty)\eta_k \, d\mu_n$$
$$\text{II} \qquad = \int (f - \|f\|_\infty)\eta_k \, d\mu_n$$

Similarly, by Fatou,

$$\liminf_{k\to\infty} \int (\|f\|_\infty - f)\eta_k \, d\mu \geq \int \|f\|_\infty - f \, d\mu$$
$$-\limsup_{k\to\infty} \int (f - \|f\|_\infty)\eta_k \, d\mu \qquad \text{III}$$

Finally, by conservation of mass,

$\int f \, d\mu$

$= \|f\|_\infty + \int f - \|f\|_\infty \, d\mu$

III

$\geq \|f\|_\infty + \limsup_{k \to \infty} \int (f - \|f\|_\infty) \eta_k \, d\mu$

II

$\geq \|f\|_\infty + \limsup_{n \to \infty} \int f - \|f\|_\infty \, d\mu_n$

$= \limsup_{n \to \infty} \int f \, d\mu_n$

$\geq \liminf_{n \to \infty} \int f \, d\mu_n$

$= -\|f\|_\infty + \liminf_{n \to \infty} \int f + \|f\|_\infty \, d\mu$

I

$\geq -\|f\|_\infty + \liminf_{k \to \infty} \int (f + \|f\|_\infty) \eta_k \, d\mu$

Fatou

$\geq -\|f\|_\infty + \int f + \|f\|_\infty \, d\mu$

$= \int f \, d\mu$

Therefore, equality holds throughout and

$$\int f\,d\mu = \lim_{n\to\infty} \int f\,d\mu_n. \qquad \square$$

Cor: If $(X,d) = (\mathbb{R}^d, |\cdot|)$, $C_c(\mathbb{R}^d)$ in $(*)$ may be replaced with $C_c^\infty(\mathbb{R}^d)$.

Pf: Suppose $(*)$ holds for all $f \in C_c^\infty(\mathbb{R}^d)$. Fix $g \in C_c(\mathbb{R}^d)$. Fix $\varepsilon > 0$, and choose $f \in C_c^\infty(\mathbb{R}^d)$ s.t. $\|f-g\|_{L^\infty} < \varepsilon$. Then

$$\left|\int g\,d\mu_n - \int g\,d\mu\right|$$

$$\leq \left|\int g\,d\mu_n - \int f\,d\mu_n\right| + \left|\int f\,d\mu_n - \int f\,d\mu\right|$$

$$+ \left|\int f\,d\mu - \int g\,d\mu\right|$$

$$< 2\varepsilon + \left|\int f\,d\mu_n - \int f\,d\mu\right|$$

Thus,

$$\limsup_{n\to\infty}\left|\int g\,d\mu_n - \int g\,d\mu\right| \leq 2\varepsilon$$

Since $\varepsilon > 0$ was arbitrary, this shows

$$\lim_{n\to\infty}\int g\,d\mu_n = \int g\,d\mu. \qquad \square$$

So we solved Kantorovich's problem...

... how does this help us solve Monge's problem?

via the Kantorovich dual problem.

Crash course in convex analysis and optimization

Let $\mathcal{X}$ be a n.v.s.

Exercise 11: Given a collection of functions $f_\alpha : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$, $\alpha \in \mathcal{A}$,

- if f_α are convex, then $\sup_\alpha f_\alpha$ is convex
- if f_α are lsc, then $\sup_\alpha f_\alpha$ is lsc

Def: Given $f: X \to \mathbb{R} \cup \{+\infty\}$ proper, its <u>conjugate</u> $f^*: X^* \to \mathbb{R} \cup \{+\infty\}$ is

$$f^*(y) = \sup_{x \in X} \{\langle y, x \rangle - f(x)\}.$$

Ex: Suppose $X = \mathbb{R}$, $f(x) = e^x$

Then $f^*(y) = \sup_{x \in \mathbb{R}} \{yx - e^x\}$

If $y < 0$, $f^*(y) = +\infty$

If $y = 0$, $f^*(y) = 0$

If $y > 0$, since $x \mapsto yx - e^x$ is a concave, differentiable fn, a critical point is a global maximizer.

$$\frac{d}{dx}(yx_* - e^{x_*}) = y - e^{x_*} = 0 \Rightarrow x_* = \log(y)$$

Then, $yx_* - e^{x_*} = y\log(y) - y$.

Thus, "entropy"

$$f^*(y) = \begin{cases} +\infty & \text{if } y < 0 \\ 0 & y = 0 \\ y\log y - y & y > 0 \end{cases}$$

Exercise 12: other examples of f^*.

Immediate consequences of the defn are...

<u>Prop</u> (Young's Inequality):
Given $f: X \to \mathbb{R} \cup \{+\infty\}$ proper,

$$f^*(y) + f(x) \geq \langle y, x \rangle \quad \forall x \in X, y \in X^*$$

Lemma: For $f: \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ proper, f^* is convex and lsc.

Pf: $f^*(y) = \sup_{\substack{x \in \mathcal{X}, \\ x \in \mathrm{dom}(f)}} \{\langle y, x \rangle - f(x)\}$.

For each $x \in \mathrm{dom}(f)$,

$$y \mapsto \langle y, x \rangle - f(x)$$

is convex and lsc.

Def: $\mathrm{dom}(f) = \{x \in \mathcal{X} : f(x) < +\infty\}$.

In a similar way, we may define

Def: Given $f: \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ s.t. f and f^* are proper, its <u>biconjugate</u> $f^{**}: \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ is

$$f^{**}(x) = \sup_{y \in \mathcal{X}^*} \{ \langle y, x \rangle - f^*(y) \}.$$

Note that, for all $x \in \mathcal{X}$, $y \in \mathcal{X}^*$,

$$f^*(y) + f(x) \geq \langle y, x \rangle$$
$$f(x) \geq \langle y, x \rangle - f^*(y)$$
$$f(x) \geq f^{**}(x)$$

Since f^{**} is <u>always</u> convex and lsc, a necessary condition for $f(x) = f^{**}(x)$ $\forall x \in \mathcal{X}$ would be that f itself is convex lsc. In fact, this is sufficient

Thm: (Fenchel-Moreau)

Given $f: X \to \mathbb{R} \cup \{+\infty\}$ proper,

(i) f is convex and lsc

$$\Updownarrow$$

f^* is proper and $f = f^{**}$

(ii) If f is convex and $f(x_0) < +\infty$,

f is lsc at $x_0 \Longleftrightarrow f(x_0) = f^{**}(x_0)$

Pf: by Hahn-Banach $\square$